Integrating fragmented evidence by network meta-analysis: relative effectiveness of psychological interventions for adults with post-traumatic stress disorder

Gerger, H ; Munder, T ; Gemperli, A ; Nüesch, E ; Trelle, S ; Jüni, P ; Barth, J

DOI: https://doi.org/10.1017/S0033291714000853

Psychological Medicine (2014), **44**, 3151–3164. © Cambridge University Press 2014
doi:10.1017/S0033291714000853

ORIGINAL ARTICLE

Integrating fragmented evidence by network meta-analysis: relative effectiveness of psychological interventions for adults with post-traumatic stress disorder

H. Gerger[1,2]*, T. Munder[1], A. Gemperli[3,4,5], E. Nüesch[1,3], S. Trelle[3], P. Jüni[1,3] and J. Barth[1,6]

[1] *Institute of Social and Preventive Medicine (ISPM), University of Bern, Switzerland*
[2] *Division of Clinical Psychology and Psychotherapy, Department of Psychology, University of Basel, Switzerland*
[3] *Department of Clinical Research, CTU Bern, University of Bern, Switzerland*
[4] *Swiss Paraplegic Research (SPF), Nottwil, Switzerland*
[5] *Department of Health Sciences and Health Policy, University of Lucerne, Switzerland*
[6] *Institute of Complementary and Integrative Medicine, University Hospital of Zurich, Switzerland*

Background. To summarize the available evidence on the effectiveness of psychological interventions for patients with post-traumatic stress disorder (PTSD).

Method. We searched bibliographic databases and reference lists of relevant systematic reviews and meta-analyses for randomized controlled trials that compared specific psychological interventions for adults with PTSD symptoms either head-to-head or against control interventions using non-specific intervention components, or against wait-list control. Two investigators independently extracted the data and assessed trial characteristics.

Results. The analyses included 4190 patients in 66 trials. An initial network meta-analysis showed large effect sizes (ESs) for all specific psychological interventions (ESs between −1.10 and −1.37) and moderate effects of psychological interventions that were used to control for non-specific intervention effects (ESs −0.58 and −0.62). ES differences between various types of specific psychological interventions were absent to small (ES differences between 0.00 and 0.27). Considerable between-trial heterogeneity occurred (τ^2=0.30). Stratified analyses revealed that trials that adhered to DSM-III/IV criteria for PTSD were associated with larger ESs. However, considerable heterogeneity remained. Heterogeneity was reduced in trials with adequate concealment of allocation and in large-sized trials. We found evidence for small-study bias.

Conclusions. Our findings show that patients with a formal diagnosis of PTSD and those with subclinical PTSD symptoms benefit from different psychological interventions. We did not identify any intervention that was consistently superior to other specific psychological interventions. However, the robustness of evidence varies considerably between different psychological interventions for PTSD, with most robust evidence for cognitive behavioral and exposure therapies.

Received 22 November 2013; Revised 24 February 2014; Accepted 11 March 2014; First published online 16 April 2014

Key words: Meta-analysis, outcome research, post-traumatic stress disorder, psychotherapy, PTSD.

Introduction

Psychological trauma is common in the general population, with lifetime prevalence between 40% and 90% (Breslau *et al.* 1991, 1998; Norris, 1992; Resnick *et al.* 1993; Kessler *et al.* 1995; Bernat *et al.* 1998). Many individuals develop psychological symptoms in the aftermath of trauma experience, commonly referred to as post-traumatic stress disorder (PTSD) symptoms. These may include re-experiencing the traumatic event, avoiding stimuli associated with the traumatic event, and increased arousal. Although many survivors of psychological trauma do not satisfy DSM-III/IV criteria for PTSD diagnosis, they may still be severely impaired and at increased risk of suicide (Stein *et al.* 1997; Marshall *et al.* 2001; Zlotnick *et al.* 2002). Trauma survivors, with or without a formal diagnosis of PTSD, often develop chronic symptoms (Kessler *et al.* 1995; Koren *et al.* 2001; Mayou *et al.* 2002; Perkonigg *et al.* 2005; Breslau, 2009) and contribute considerably to health-care costs (Walker *et al.* 1999).

A variety of psychological interventions have been suggested to treat PTSD symptoms. Some are based

* Address for correspondence: H. Gerger, Ph.D., Department of Clinical Psychology and Psychotherapy, University of Basel, Missionsstrasse 62, 4055 Basel, Switzerland.
(Email: heike.gerger@gmail.com)

on aetiological models and propose specific intervention components to relieve symptoms, such as exposure to trauma-related stimuli (Foa *et al.* 1991) or working through cognitions associated with the trauma (Ehlers *et al.* 2005). Other psychological interventions are based on components that are not unique to any particular psychological intervention and are therefore often described as 'common' or 'non-specific', for example empathy in supportive therapies or attention effects in relaxation treatments. Several standard meta-analyses have been designed to determine which psychological interventions are most promising for patients with PTSD symptoms, but they have been unable to come to a definite conclusion (Bradley *et al.* 2005; Benish *et al.* 2008; Bisson & Andrew, 2007).

There are two possible sources of evidence for a comparison of the effectiveness of two interventions A and B. The first is a direct within-trial comparison of intervention A and B (direct evidence). The second is a comparison of results from trials that compare either of the two interventions A and B with a common third intervention C (indirect evidence). Although direct, within-trial comparisons of active psychological interventions are the gold standard for establishing relative effectiveness, for many psychological interventions direct within-trial comparisons are rare or even non-existent. Therefore, both sources of evidence are useful for a comprehensive evaluation of relative effectiveness of interventions A and B.

Two of the above-mentioned meta-analyses took two approaches to determining the relative effects of different psychological interventions. One research group (Bisson & Andrew, 2007) conducted separate meta-analyses for each direct comparison of psychological interventions that occurred in the literature (direct evidence). They included trials that compared the different interventions head-to-head, or with control interventions, and compared effects across single meta-analyses (informal indirect evidence). When effects of single meta-analyses were equal, they assumed that those interventions were equally effective even when interventions had not been directly compared within a single trial. Heterogeneity remained unexplained in some cases, and this limited interpretation of their results. By this method, they concluded that trauma-focused cognitive behavioral therapy (CBT) and eye movement desensitization and reprocessing (EMDR) were more effective than other psychological interventions, including supportive therapy (ST) and psychodynamic therapy (PT) for instance.

A second group (Benish *et al.* 2008) included only trials that compared different types of psychological interventions head-to-head (direct evidence) and pooled them in one meta-analysis. They restricted the range of psychological interventions to those they classified as 'intended to be therapeutic'. This classification included psychological interventions such as PT, but excluded ST or stress management (SM). The latter interventions are typically used to control for non-specific intervention effects and the authors therefore did not consider them to be 'intended to be therapeutic'. The authors interpreted the absence of between-trial heterogeneity of effect sizes as an indicator that different psychological interventions were equally effective and concluded that all interventions that were 'intended to be therapeutic' were equally effective.

Other meta-analyses that considered only direct evidence for the effectiveness of particular types of psychological interventions found no difference in the effectiveness of EMDR and trauma-focused CBT (Seidler & Wagner, 2006), EMDR and exposure-based therapy (ET) (Davidson & Parker, 2001), CBT, ET and cognitive therapy (CT) (Mendes *et al.* 2008) and ET, CBT, EMDR and CT (Powers *et al.* 2010).

The results of recent meta-analyses leave us with a patchwork of findings based on direct evidence and informal indirect evidence. Although the equivalent effectiveness of trauma-focused CBT, ET, CT and EMDR seems to have been established, the effectiveness of other psychological interventions (including PT, ST and SM) has not yet been ascertained. Residual heterogeneity complicates the interpretation of previous findings in many cases (Davidson & Parker, 2001; Bradley *et al.* 2005; Bisson & Andrew, 2007; Powers *et al.* 2010).

To overcome the limitations of the available comparisons and informal indirect evidence, we used network meta-analysis, a methodological approach to integrate trials, comparing a variety of psychological interventions head-to-head or with a control condition (Lumley, 2002; Lu & Ades, 2004; Salanti *et al.* 2008; Cipriani *et al.* 2009). In network meta-analyses, the information available from within-trial comparisons of interventions A and B is combined with indirect comparisons of A and B derived from trials that compare either of the two interventions with a common comparator C (either a third psychological intervention or a control condition). Network meta-analysis has previously been used to investigate the effectiveness of pharmacological treatments for depression (Cipriani *et al.* 2009) and mania (Cipriani *et al.* 2011), and in the evaluation of psychological interventions for depression (Barth *et al.* 2013).

Meta-analyses on psychological interventions for PTSD used different approaches to classify such interventions. As a result, the number and definition of categories vary across meta-analyses. Some researchers used a large number of categories to capture differences between interventions according to their

theoretical backgrounds (e.g. differentiating between mainly cognitive interventions, primarily exposure-based interventions and a mixture of cognitive and behavioral elements; Bradley *et al.* 2005; Mendes *et al.* 2008; Watts *et al.* 2013). Other authors suggested summarizing interventions with different theoretical backgrounds to broad categories (e.g. trauma-focused CBT including CBT, CT and ET; Bisson & Andrew, 2007). Finally, some authors even argued against categorizing interventions at all (Benish *et al.* 2008; Wampold *et al.* 2010) because differences in their effectiveness are only small. We chose an approach that allowed us to include a large number of direct comparisons between psychological interventions in the network (e.g. from dismantling studies such as Resick *et al.* 2008). At the same time we aimed at limiting the number of categories to a reasonable number (e.g. no differentiation between different types of exposure such as *in vivo* and *in sensu*).

To examine whether different approaches to classifying psychological interventions affect meta-analytic results, we reduced the number of intervention categories subsequently and looked at possible changes in effect sizes, heterogeneity statistics and model fit.

As the quality of primary studies is known to be a potential threat to the validity of meta-analyses (Jüni *et al.* 2001; Matt & Cook, 2009; Cuijpers *et al.* 2010*b*), we assessed the influence of study quality (Wood *et al.* 2008; Nüesch *et al.* 2009), sample size (Nüesch *et al.* 2010) and type of outcome assessment (Cuijpers *et al.* 2010*a*). We also controlled for the presence of a formal PTSD diagnosis in our analyses.

Method

Literature search

The literature search was based on a comprehensive initiative to build a database of references to clinical trials that had investigated the effectiveness of any psychological intervention for adults with PTSD symptoms. We searched bibliographic databases relevant for the field of psychotherapy (EMBASE, Medline, PsycINFO, Cochrane Controlled Trials Register and PSYNDEX) by combining key words and text words related to psychological interventions, randomized trials and PTSD (see online Appendix 1 for the search strategies used). We also checked the reference lists of relevant systematic reviews and meta-analyses (van Etten & Taylor, 1998; Bradley *et al.* 2005; Benish *et al.* 2008; Bisson & Andrew, 2008; Cloitre, 2009). The search was performed in January 2011 for trials published between 1980 and 2010.

Selection of trials

We included randomized trials in adults with full or subclinical PTSD that compared specific psychological interventions head-to-head (e.g. CBT compared with ET) against wait-list (WL), or against another control intervention using only non-specific intervention components such as therapist alliance, general attention or empathy (e.g. ST). Other potential control interventions, such as standard care involving pharmacological intervention or the use of pill placebos, were not eligible. Patients were considered to have subclinical PTSD if they had experienced at least one psychological trauma according to DSM-IV criteria and reported subsequent PTSD symptoms. We included both veteran and civilian samples. For a specific psychological intervention to qualify, it had to be implemented at the individual level (rather than as group, family or couples therapy), include face-to-face contact between the patient and the therapist (as opposed to telephone or internet-based interaction between patient and therapist), consist mainly of verbal communication, and directly address the trauma or subsequent PTSD symptoms. Trials had to be published as full journal articles; there were no language restrictions. We contacted the authors if the available information was not sufficient to determine inclusion of the trial. Seven investigators (H.G., T.M., one Ph.D. student and four M.Sc. students) determined eligibility according to a structured manual. Eligibility of a random sample of 200 references was determined by all seven investigators; the κ statistic for the coding of a reference as clearly included, clearly excluded or unclear based on the title and the abstract was 0.73. There were no disagreements about whether a reference could be excluded based on title and abstract. In ambiguous cases, a decision was made by consensus between H.G., T.M. and a senior researcher (J.B.) based on the full text.

Outcome measures

The prespecified primary outcome was severity of PTSD symptoms after the intervention or at maximum of 1 month after the intervention was terminated, measured with a validated scale. We preferred data from scales that assessed symptoms according to DSM-III/IV diagnostic criteria over (sub)scales that focused on only one symptom cluster. When more than one outcome measure was reported, we extracted the highest outcome on a predefined hierarchy. Most frequently used scales were given precedence. Self-rated PTSD symptoms were preferred to observer-rated PTSD symptoms, and results from intention-to-treat (ITT) analyses that included all randomized patients took precedence over results from analyses that excluded patients. Both self-rated outcome assessment and ITT

analyses have been shown to result in conservative effect estimates (Nüesch *et al.* 2009; Cuijpers *et al.* 2010*a*).

Data extraction and coding

We classified interventions according to eight prespecified categories: WL, SM, ST, ET, CT, EMDR, CBT, and other psychological interventions (OPIs; see online Appendix 2 for descriptions of interventions). This classification was based primarily on the treatment descriptions in the published study reports. For further analyses we combined single categories. First, we reduced the number of intervention categories that relied on cognitive behavioral components: from three individual categories (i.e. CBT, CT and ET) to two (i.e. CBT_c and ET), to one broad CBT category (CBT_b); EMDR, OPI, ST, SM and WL remained unchanged. Second, we further reduced the number of categories to three, with all interventions that were based on specific intervention components for PTSD in one category of specific psychological interventions (i.e. CBT, EMDR, CT, ET and OPI), and interventions that were used as control for non-specific intervention components that are common to all psychological interventions as the second category (i.e. ST and SM), and WL as the third category.

Studies were classified according to their adherence to DSM-III/IV criteria for PTSD during patient inclusion. Studies in which at least 80% of patients satisfied DSM-III/IV criteria for PTSD were considered to adhere to DSM-III/IV criteria for PTSD during patient inclusion.

We assessed concealment of treatment allocation (Jüni *et al.* 2001; Wood *et al.* 2008), the reporting of ITT data (Nüesch *et al.* 2009) and the type of outcome assessment (Cuijpers *et al.* 2010*a*). Concealment of allocation was considered adequate if the investigators responsible for patient selection did not suspect which treatment was next before allocation. Analyses were considered adequate if all recruited patients were analysed in the group to which they were originally allocated, regardless of intervention received (ITT principle). Analyses were considered inadequate if data were insufficient to calculate ES based on the ITT sample. Outcome assessment was classified as self-rated when the patients used self-rating scales for outcome assessment, and as observer rated when some other person was involved in data collection, for example through clinical interviews.

When necessary, means and measures of dispersion of clinical outcome data were approximated from figures in the reports. All trial data were extracted in duplicate on a standardized form (Epidata 3.1, The Epidata Association, Denmark) by two out of five investigators (H.G. or T.M. and three M.Sc. students). All investigators were trained with a manual in a 2-day training workshop. Disagreements were resolved after they had been reviewed by a third investigator. The median κ across all extracted clinical and methodological characteristics was 0.79 (range 0.62–0.97).

Statistical analysis

For each treatment arm, we standardized mean values at the end of treatment using the pooled standard deviation (S.D.) across arms within each trial. If S.D.s were not provided, we calculated them from standard errors (S.E.s), confidence intervals (CIs) or other measures as described elsewhere (Follmann *et al.* 1992; Reichenbach *et al.* 2007).

For the network meta-analysis, we used an extension of Bayesian random effects models for comparisons of multiple interventions (Smith *et al.* 1995; Lu & Ades, 2004). It considers all included comparisons between interventions, while completely preserving randomization within each trial and accounting for correlation between multiple comparisons within a trial with more than two treatment arms (Cooper *et al.* 2006). Pooled effect sizes (ESs) were derived by the median of the posterior distribution of the difference in standardized mean values of two treatments. Negative ESs indicate the experimental intervention had a beneficial effect and may be interpreted as described elsewhere (Cohen, 1988), with −0.20 S.D. units representing a small, −0.50 a moderate and −0.80 a large difference between interventions. Corresponding 95% credibility intervals (CrIs) were derived by the 2.5th and 97.5th percentiles of the posterior distribution. The between-trial heterogeneity estimate τ^2 was also estimated from the median of the corresponding posterior distribution. τ represents the S.D. of the underlying distribution from which the included trials are assumed to be a random sample. Based on our definition of small, moderate and large differences between interventions, we interpreted τ^2 as follows: $\tau^2=0.01$ $[(0.2/2)^2]$ was considered to represent low heterogeneity, $\tau^2=0.0625$ $[(0.5/2)^2]$ moderate heterogeneity and $\tau^2=0.16$ $[(0.8/2)^2]$ high heterogeneity between studies. τ^2 has been shown to be independent of the number of studies and the number of patients included in the meta-analysis (i.e. no increase with large numbers of studies or large sample sizes; Rücker *et al.* 2008). The consistency of the network was determined by comparing effect estimates derived by a meta-analysis including only direct comparisons with the indirect effect estimates derived by a network meta-analysis excluding the respective direct comparison. This procedure was applied to all existing pair-wise comparisons in the analysis dataset. Goodness-of-fit of the model was assessed with *Q*–*Q* plots.

To determine whether the network of evidence was affected by small-study effects, we drew contour-enhanced funnel plots (Peters *et al.* 2008) and added lines representing predicted intervention effects derived from random effects meta-regression using the S.E. as the explanatory variable. Then we assessed funnel plot asymmetry with a regression test (Egger *et al.* 1997).

To determine whether estimated intervention effects were affected by trial characteristics, we performed stratified analyses by including an interaction term of treatment and trial characteristics as covariates in the network meta-analysis. We considered the following characteristics: adequate concealment of allocation, ITT analysis performed, trial size, type of outcome assessment and adherence to DSM-III/IV criteria for PTSD. p values for interaction effects between trial characteristics and intervention effects were estimated from the posterior distribution of covariates. These p values can be interpreted in the same way as traditional p values for interaction (Altman & Bland, 2003). We used two cut-offs for trial size. The first was based on the median of 19 patients per trial arm observed in included studies, and distinguished between very small trials with an average of 19 patients or less per arm, and trials with 20 patients or more. The second cut-off distinguished between small to moderate trials with an average of 59 patients or less per arm, and trials with 60 patients or more. This trial size yields more than 90% power to detect a moderate to large ES of −0.60 S.D. units at a two-sided $\alpha=0.05$.

Finally, we investigated whether different approaches to classifying psychological interventions affect meta-analytic results. We subsequently combined single interventions into broader categories. This was implemented through a fully Bayesian strategy including the network with all psychological interventions, but effect parameters were restricted to be equal for comparisons of interventions between groups (e.g. the specific and the non-specific psychological interventions) and zero for comparison of interventions within the same group (e.g. interventions within the CBT_b category). The network with effects for every psychological intervention was compared with a network based on the group-wise effects through goodness-of-fit using the deviance information criterion (DIC; Spiegelhalter *et al.* 2002). We used Stata releases 11 and 12 (StataCorp LP 2005, USA) and WinBUGS version 1.4.3 (MRC Biostatistics Unit 2007, UK) for all analyses.

Results

Initially, we identified 1311 references in our literature search and found 341 to be potentially eligible (Fig. 1). Sixty-six trials met our criteria and were included in the initial network meta-analysis (see online Appendix 7 for references and online Appendix 3 for a detailed description of each trial). The 66 trials had 155 arms that qualified for the analysis, with a median of 19 patients per arm (range 6–143), and a total of 4190 randomized patients. All but one trial was published in English. Thirty-two trials were conducted in the USA (48%). The median year of publication was 2003 (range 1989–2010). The most frequently evaluated specific psychological interventions were CBT in 31 trials (47%), ET in 23 trials (35%), EMDR in 20 trials (30%), followed by OPI in 14 trials (21%) and CT in six trials (9%). As control groups, WL was used in 37 trials (56%), followed by ST in 11 trials (17%) and SM in seven trials (11%) as non-specific control interventions (see Fig. 2 for the network of evidence).

Initial network meta-analysis

All 66 trials contributed to the overall network meta-analysis. Table 1 presents ESs of all interventions compared with WL. The five interventions, all based on specific psychological components, were associated with large ESs between −1.10 and −1.37. The ESs of ST and SM as non-specific control interventions were moderate (ES=−0.62 and −0.58 respectively). However, the only significant difference between two psychological interventions was that EMDR outperformed ST. Figure 3*a* presents an overview of pair-wise comparisons (ESs with 95% CrIs) of all interventions. The τ^2 estimate of 0.30 suggested very large heterogeneity. There was no evidence of network inconsistency: although direct and indirect effect estimates differed in a range of moderate to large ESs for some comparisons, all CIs overlapped zero (see online Appendix 5).

Exploration of variation between trials

Network meta-analyses that were stratified according to different characteristics of the included trials showed that the adherence to DSM-III/IV criteria for PTSD during patient inclusion was the most relevant moderator, with $p=0.01$ (Table 1). ESs were larger in trials that adhered to DSM-III/IV during patient inclusion (with the smallest ES of −0.76 for SM and the largest ES of −1.58 for EMDR) and ESs were smaller in trials that included patients with subclinical PTSD (with the smallest ES of −0.10 for ST and the largest ES of −0.87 for CBT).

Examination of the funnel plot of all psychological interventions compared with WL indicated asymmetry with missing trials in areas of non-significance, even though the corresponding regression test was only borderline positive ($p=0.053$; online Appendix 6).

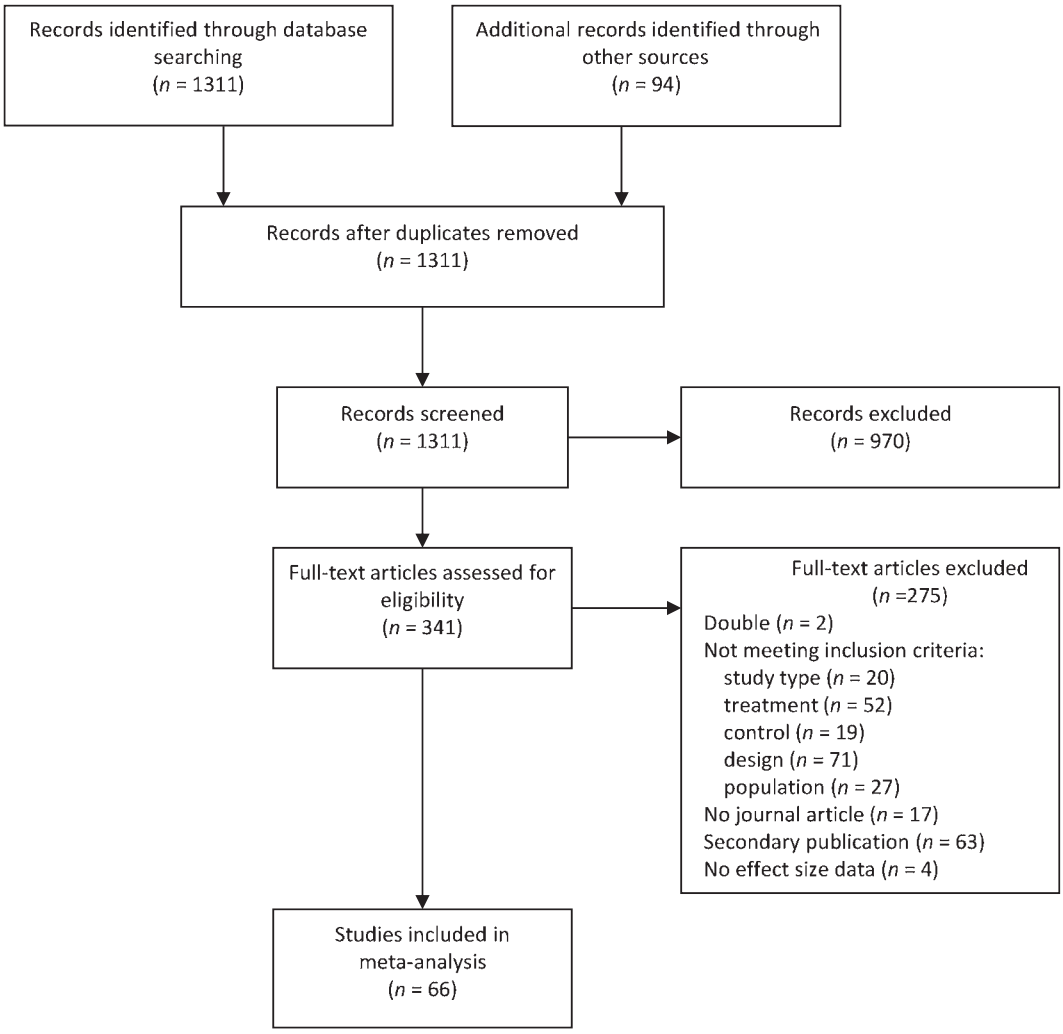

Fig. 1. Flowchart of trial selection.

Heterogeneity remained high in all strata except for trials with adequate concealment of allocation and trials with a large trial size. A large amount of heterogeneity between effect estimates from individual trials complicated the interpretation of results in most subgroups of trials. Therefore, we highlight the results of the two subgroups of studies with moderate heterogeneity, which allow clearer conclusions to be drawn.

Trials with adequate concealment of allocation

In the 10 trials with adequate concealment of allocation, CBT was used as the specific psychological intervention in five trials, EMDR in one trial, CT in four trials, ET in four trials and OPI in two trials. WL was used as control intervention in four trials and ST in one trial. The test of interaction was non-significant (p=0.98). The five interventions, all based on specific psychological components, had large ESs. The ES of ST as a non-specific control intervention was moderate. The only significant difference between two psychological interventions was that EMDR outperformed ST (see online Appendix 4). Table 1 shows that the differences between ESs in the adequately *versus* inadequately concealed trials were only small. Between-trial heterogeneity was moderate (τ^2=0.04).

Large-sized trials

In the seven large-sized trials, CBT was used as the specific psychological intervention in five trials, ET in three trials, and traumatic incident reduction (TIR) therapy classified as OPI in one trial. WL was used as control intervention in five trials and ST in two trials. ESs of CBT and ET compared with WL were smaller than in the initial network meta-analysis (ES=–0.86 and −0.75 respectively) whereas the ES of ST as one of the non-specific control interventions did not change (ES=–0.61). However, the lower

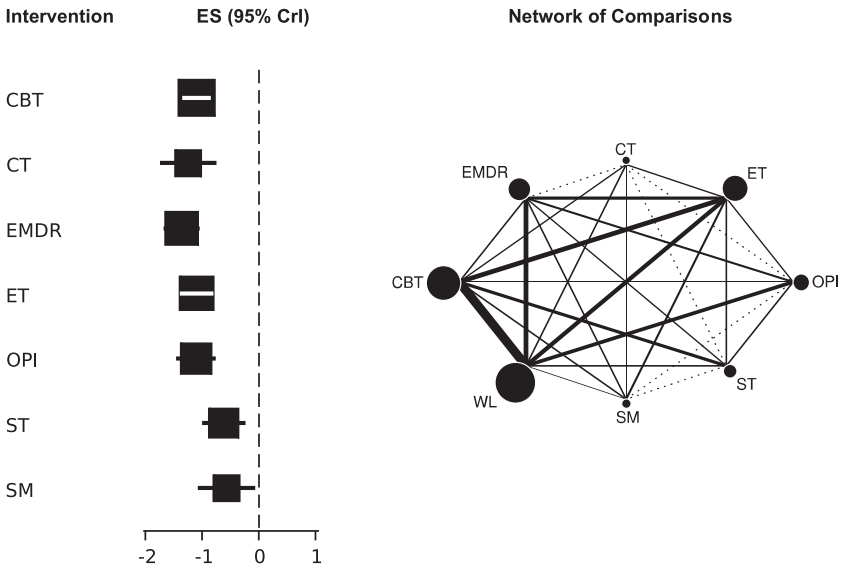

Fig. 2. Effect sizes (ESs) and credibility intervals (CrIs) of all interventions compared with wait-list (WL) in the initial network meta-analysis including 66 trials and network of eligible comparisons between different psychological interventions. Negative ESs indicate superiority of the intervention compared to WL. Dotted lines indicate which comparisons have not yet been evaluated in a randomized controlled trial. The width of the lines is proportional to the number of trials comparing each pair of interventions, and the size of each node is proportional to the number of trials. CBT, Cognitive behavioral therapy; CT, cognitive therapy; EMDR, eye movement desensitization and reprocessing; ET, exposure therapy; OPI, other psychological intervention; SM, stress management; ST, supportive therapies.

number of included trials reduced the precision of the estimates and the ES for ST was no longer statistically significant. For TIR therapy classified as OPI, the benefit was smaller (ES=–0.37) but 95% CrIs of all psychological interventions under investigation were largely overlapping: pair-wise comparisons showed that none of the three specific psychological interventions was superior to ST (online Appendix 4). Between-trial heterogeneity was moderate (τ^2=0.08).

Reduction of the number of intervention categories

Table 2 shows the results from different models in which the number of intervention categories was reduced subsequently. Combining single categories into broader categories did not affect either ESs or heterogeneity statistics. In all models we found no superiority of any specific psychological intervention over any other specific psychological intervention (Fig. 3*b*, *c*). However, we found evidence for the superiority of specific over non-specific psychological interventions. When all specific interventions were summarized into one broad category and compared to the non-specific psychological interventions, we found a moderate superiority of −0.55 (95% CrI −0.83 to −0.28, τ^2=0.29) of specific over non-specific psychological interventions. The goodness-of-fit of either modelling all interventions separately or using the grouping was comparable (DICs between 102.53 and 104.72) with a slight decrease in the models with fewer categories (Table 2).

Discussion

Main findings

Our network meta-analysis of different psychological interventions in patients with PTSD symptoms, which integrated direct and indirect evidence, suggests that different specific psychological interventions for treating PTSD have similar benefits. We found no evidence for the benefit of differentiating between different types of specific psychological interventions with respect to their effectiveness. We found evidence for the presence of small-study bias (i.e. the overestimation of intervention effects in trials with small to moderate trial size). Effect estimates were based on robust evidence with respect to a reasonable number of trials and the size of the trials only for psychological interventions that are based on cognitive behavioral intervention components. Our results from large-sized trials, however, indicate that STs that are based on non-specific intervention components might be equally effective as psychological interventions developed specifically for the treatment of PTSD symptoms (i.e. CBT and ET). It is important to note, however, that

Table 1. *PTSD outcomes of all interventions compared to wait-list stratified according to methodological and clinical trial characteristics*

	No. of trials (patients)	SM	ST	OPI	ET	EMDR	CT	CBT	τ^2	p^a
		ES (95% CrI)	ES (95% CrI)	ES (95% CrI)	ES (95% CrI)	ES (95% CrI)	ES (95% CrI)	ES (95% CrI)		
All trials	66 (4190)	**−0.58 (−1.09 to −0.06)**	**−0.62 (−1.01 to −0.23)**	**−1.11 (−1.47 to −0.76)**	**−1.10 (−1.40 to −0.81)**	**−1.37 (−1.69 to −1.05)**	**−1.25 (−1.75 to −0.76)**	**−1.10 (−1.36 to −0.85)**	0.30	
Adequate concealment										0.78
Yes	10 (996)	N.E.	−0.46 (−1.37 to 0.43)	**−1.02 (−1.78 to −0.24)**	**−0.84 (−1.49 to −0.22)**	**−1.65 (−2.36 to −0.91)**	**−1.03 (−1.50 to −0.57)**	**−1.05 (−1.52 to −0.55)**	0.04	
No or unclear	56 (3194)	**−0.60 (−1.16 to −0.05)**	**−0.61 (−1.06 to −0.17)**	**−1.11 (−1.52 to −0.70)**	**−1.13 (−1.48 to −0.79)**	**−1.36 (−1.71 to −1.00)**	**−1.51 (−2.41 to −0.63)**	**−1.12 (−1.41 to −0.82)**	0.37	
ITT analyses										0.47
Yes	21 (1986)	N.E.	**−0.76 (−1.32 to −0.21)**	**−1.11 (−1.83 to −0.39)**	**−0.95 (−1.42 to −0.49)**	**−3.01 (−4.12 to −1.92)**	**−1.03 (−1.80 to −0.25)**	**−1.07 (−1.43 to −0.72)**	0.21	
No or unclear	45 (2204)	**−0.56 (−1.07 to −0.05)**	−0.37 (−0.89 to 0.15)	**−1.04 (−1.44 to −0.66)**	**−1.16 (−1.53 to −0.79)**	**−1.25 (−1.58 to −0.92)**	**−1.36 (−1.97 to −0.77)**	**−1.04 (−1.38 to −0.71)**	0.28	
Trial size										0.15
Moderate to large (⩾20)	33 (3226)	**−0.77 (−1.45 to −0.08)**	**−0.68 (−1.09 to −0.28)**	**−0.77 (−1.15 to −0.39)**	**−1.10 (−1.41 to −0.79)**	**−1.42 (−1.82 to −1.03)**	**−1.07 (−1.52 to −0.63)**	**−0.96 (−1.21 to −0.71)**	0.17	
Small (<20)	33 (964)	−0.51 (−1.27 to 0.26)	−0.59 (−1.31 to 0.15)	**−1.54 (−2.20 to −0.90)**	**−0.99 (−1.53 to −0.44)**	**−1.45 (−1.96 to −0.93)**	**−2.24 (−3.81 to −0.68)**	**−1.33 (−1.84 to −0.82)**	0.46	
Trial size										0.16
Large (⩾60)	7 (1233)	N.E.	−0.61 (−1.37 to 0.15)	−0.37 (−1.20 to 0.48)	**−0.75 (−1.35 to −0.12)**	N.E.	N.E.	**−0.86 (−1.33 to −0.37)**	0.08	
Small to moderate (<60)	59 (2957)	**−0.63 (−1.15 to −0.09)**	**−0.61 (−1.05 to −0.17)**	**−1.21 (−1.60 to −0.82)**	**−1.18 (−1.50 to −0.86)**	**−1.41 (−1.74 to −1.08)**	**−1.29 (−1.81 to −0.77)**	**−1.16 (−1.44 to −0.87)**	0.33	
Diagnosis										**0.01**
PTSD	47 (3032)	**−0.76 (−1.34 to −0.18)**	**−0.97 (−1.52 to −0.42)**	**−1.11 (−1.66 to −0.57)**	**−1.25 (−1.59 to −0.92)**	**−1.58 (−1.99 to −1.17)**	**−1.35 (−1.88 to −0.82)**	**−1.29 (−1.61 to −0.96)**	0.33	
Subclinical PTSD	19 (1158)	0.29 (−0.99 to 1.58)	−0.10 (−0.60 to 0.43)	**−0.84 (−1.29 to −0.43)**	−0.26 (−1.37 to 0.86)	**−0.85 (−1.32 to −0.35)**	N.E.	**−0.87 (−1.24 to −0.51)**	0.14	
Outcome assessment										0.58
Self	54 (3298)	−0.45 (−1.08 to 0.18)	**−0.71 (−1.16 to −0.25)**	**−1.03 (−1.44 to −0.62)**	**−1.02 (−1.37 to −0.67)**	**−1.42 (−1.75 to −1.08)**	**−1.31 (−1.88 to −0.76)**	**−1.19 (−1.48 to −0.89)**	0.30	
Observer	12 (892)	−0.55 (−1.63 to 0.48)	−0.50 (−1.34 to 0.42)	**−1.30 (−2.13 to −0.50)**	**−1.20 (−1.90 to −0.55)**	−0.46 (−1.90 to 0.91)	−0.96 (−2.22 to 0.27)	**−0.80 (−1.35 to −0.25)**	0.28	

PTSD, Post-traumatic stress disorder; CBT, cognitive behavioral therapy; CT, cognitive therapy; CrI, credibility interval; EMDR, eye movement desensitization and reprocessing; ES, effect size; ET, exposure therapy; ITT, intention-to-treat; OPI, other psychological intervention; τ^2, variability between trials; SM, stress management; ST, supportive therapies; N.E., not estimated: if CrIs were larger than 20 standard deviation (S.D.) units.

[a] The p value indicates whether the difference between subgroups is significant.

Bold font indicates whether the ES was statistically significant.

(*a*) Eight intervention categories: Initial classification (*N*=66)

SM						
-0.03 -0.64, 0.54	ST					
-0.53 -1.13, 0.06	**-0.49** **-0.94, -0.05**	OPI				
-0.52 **-1.04, -0.02**	**-0.48** **-0.90, -0.08**	0.00 -0.40, 0.41	ET			
-0.79 **-1.33, -0.28**	**-0.75** **-1.19, -0.33**	-0.27 -0.66, 0.15	-0.27 -0.62, 0.08	EMDR		
-0.67 **-1.33, -0.02**	**-0.63** **-1.23, -0.03**	-0.14 -0.72, 0.44	-0.15 -0.66, 0.37	0.12 -0.43,0.69	CT	
-0.52 **-1.03, -0.04**	**-0.48** **-0.87, -0.11**	0.01 -0.38, 0.40	0.00 -0.30, 0.30	0.27 -0.07, 0.60	0.15 -0.36, 0.65	CBT

(*b*) Seven intervention categories: Combined CBT and CT (*N*=66)

SM					
-0.08 -0.67, 0.53	ST				
-0.57 -1.15, 0.02	**-0.49** **-0.93, -0.04**	OPI			
-0.55 **-1.05, -0.03**	**-0.47** **-0.88, -0.06**	0.02 -0.38, 0.43	ET		
-0.83 **-1.35, -0.28**	**-0.74** **-1.18, -0.31**	-0.25 -0.66, 0.16	-0.27 -0.63, 0.07	EMDR	
-0.58 **-1.08, -0.07**	**-0.50** **-0.88, -0.12**	-0.01 -0.39, 0.37	-0.03 -0.32, 0.26	0.25 -0.10, 0.58	CBTc

(*c*) Six intervention categories: Combined CBT, CT and ET (*N*=66)

SM				
-0.08 -0.66, 0.52	ST			
-0.56 -1.15, 0.03	**-0.48** **-0.94, -0.05**	OPI		
-0.82 **-1.35, -0.29**	**-0.74** **-1.18, -0.33**	-0.26 -0.67, 0.14	EMDR	
-0.57 **-1.04, -0.08**	**-0.49** **-0.84, -0.13**	-0.00 -0.37, 0.37	0.26 -0.06, 0.57	CBTb

Fig. 3. Effect sizes (ESs) and credibility intervals (CrIs) of pair-wise comparisons based on three different models of classifying interventions. Negative ESs indicate superiority of the row defining intervention compared with the column defining intervention. Bold font indicates whether the ES was statistically significant. CBT, Cognitive behavioral therapy; CBT_b, broad CBT category; CBT_c, CBT with focus on cognitions; CT, cognitive therapy; EMDR, eye movement desensitization and reprocessing; ET, exposure therapy; OPI, other psychological intervention; SM, stress management; ST, supportive therapy; WL, wait-list.

interventions in the ST category may vary considerably in their actual content. Therefore, the conclusion that ST, CBT and ET may be equally effective for patients with PTSD symptoms seems premature and deserves more examination.

Strengths and weaknesses of the study

This study has several strengths. We performed an extensive literature search (Egger *et al.* 2003). To minimize bias and transcription errors, data extraction

Table 2. *Results from different models with combined intervention categories: ESs, heterogeneity and model fit*

Model	ES compared with WL (95% CrI)	τ^2	DIC
All interventions: initial classification		0.30	104.72
CBT	−1.10 (−1.36 to −0.85)		
CT	−1.25 (−1.75 to −0.76)		
EMDR	−1.37 (−1.69 to −1.05)		
ET	−1.10 (−1.40 to −0.81)		
OPI	−1.11 (−1.47 to −0.76)		
ST	−0.62 (−1.01 to −0.23)		
SM	−0.58 (−1.09 to −0.06)		
Seven intervention categories: combined CBT, CT		0.31	103.15
CBT_c	−1.13 (−1.38 to −0.89)		
EMDR	−1.38 (−1.69 to −1.05)		
ET	−1.10 (−1.40 to −0.80)		
OPI	−1.12 (−1.48 to −0.76)		
ST	−0.63 (−1.02 to −0.25)		
SM	−0.55 (−1.09 to −0.04)		
Six intervention categories: combined CBT, CT, ET		0.30	103.46
CBT_b	−1.12 (−1.34 to −0.90)		
EMDR	−1.38 (−1.70 to −1.07)		
OPI	−1.11 (−1.47 to −0.77)		
ST	−0.63 (−1.02 to −0.26)		
SM	−0.55 (−1.07 to −0.05)		
Three intervention categories: combined CBT, CT, EMDR, ET, OPI and ST, SM		0.29	102.53
Specific	−1.16 (−1.36 to −0.98)		
Non-specific	−0.61 (−0.94 to −0.28)		

CBT, Cognitive behavioral therapy; CBT_b, broad CBT category; CBT_c, CBT with focus on cognitions; CrI, credibility interval; CT, cognitive therapy; DIC, goodness-of-fit using the deviance information criterion; EMDR, eye movement desensitization and reprocessing; ES, effect size; ET, exposure therapy; OPI, other psychological intervention; SM, stress management; ST, supportive therapies; τ^2, variability between trials; WL, wait-list.

was performed electronically and independently by two investigators (Egger *et al.* 2001). Components used for quality assessment are validated and reported to be associated with bias (Jüni *et al.* 2001; Wood *et al.* 2008; Nüesch *et al.* 2009). Our network meta-analysis integrated all available evidence on the effectiveness of psychotherapy from direct and indirect comparisons into one analysis while fully preserving randomization.

This research has some limitations. Like standard meta-analysis, network meta-analysis assumes the included trials are drawn from the same population. This assumption implies, first, that heterogeneity between ESs of individual trials is small, and second, that direct and indirect effect estimates do not differ significantly (i.e. no inconsistency between directly and indirectly estimated ESs). If present, both heterogeneity and inconsistency complicate the interpretation of results. Although we found no evidence for inconsistency in any analysis, heterogeneity was reduced to a moderate amount only in some subsets of trials (i.e. trials with adequate concealment of allocation and large-sized trials).

We included only published trials in our analysis. In view of the skewed funnel plot, we suggest that including unpublished material would probably have resulted in smaller estimated benefits than those observed in our study.

We did not analyse the possibly moderating effect of additional characteristics of the patient sample, such as chronicity of the symptoms, the type of trauma or whether it was a veteran or civilian sample. There is some evidence that such characteristics moderate relative effects between specific and non-specific psychological interventions for PTSD (Gerger *et al.* 2013). Nor did our study control for researcher allegiance bias (Munder *et al.* 2012), which may have introduced bias in effect estimates. Because data on a comparison level such as allegiance cannot be considered in network meta-analysis, it is likely that researcher

preferences influenced the intervention effects found in this study to some extent.

Finally, the association of ESs with the adherence to DSM criteria for PTSD during patient inclusion is based on aggregated data available on the level of trials rather than on individual patient data. This approach is susceptible to the ecological fallacy (Thompson & Higgins, 2002).

Comparison with other studies

Our findings of large ESs of specific psychological interventions and moderate ESs of non-specific psychological interventions are in line with findings from previous meta-analyses (Bradley *et al.* 2005; Bisson & Andrew, 2007). Our results confirmed previous findings that the effects of psychological interventions for PTSD are likely to be overestimated (Watts *et al.* 2013). However, our results extend previous findings because our analyses suggest that such small-study bias was present particularly in specific psychological interventions for PTSD but not in non-specific psychological interventions. This finding is based on exploratory analyses, however, and needs further examination. The finding that STs may be as effective as CBT and ET from our analysis of large-sized trials is confirmed by results from a recent meta-analysis on present-centred therapy, one of the STs in our meta-analysis (Frost *et al.* 2014). The authors found only a small and non-significant ES difference between present-centred therapies and specific psychological interventions for PTSD on PTSD symptoms.

Classifying interventions into mutually exclusive categories is a common procedure in meta-analyses that examine the effectiveness of different types of interventions. In previous meta-analyses we found various numbers of categories in addition to variation in the labelling of individual categories (e.g. Bradley *et al.* 2005; Bisson & Andrew, 2007; Watts *et al.* 2013). To reduce between-trial heterogeneity, Watts *et al.* (2013) subdivided their initial classification into more specific intervention types subsequently. Such an approach has the disadvantage that it results in a large number of intervention categories. As a consequence, the number of single meta-analyses that need to be conducted increases while the number of trials summarized in one meta-analysis decreases. By contrast, our network meta-analyses included all trials at the same time. We started with a larger number of categories and reduced the number of categories (i.e. number of nodes in the network) in subsequent analyses. We found no evidence that classifying specific psychological interventions into single categories according to their underlying theoretical background would reduce heterogeneity or provide a better model fit. This finding can be seen as confirmation of the conclusion drawn by Benish *et al.* (2008) and Wampold *et al.* (2010), who deny the worth of classifying specific psychological interventions into categories. For the initial classification of psychological interventions, we adhered to the labels given by the authors of primary studies in most cases. However, we cannot rule out the possibility that interventions within the same category differed as to their content, which might have contributed to heterogeneity.

The association of intervention benefits with methodological quality varied for different types of psychological interventions. This inconclusive pattern is in line with previous investigations (Bradley *et al.* 2005; Bisson & Andrew, 2007). The results from our analyses are difficult to interpret because a substantial amount of heterogeneity remained unexplained in the stratified analyses. This may be related to the generally unsatisfactory methodological quality and the predominantly small sample size of included trials. Besides this, we calculated interaction effects and the corresponding *p* values based on a model that assumes the same interaction effect for all comparisons. A more flexible modelling approach would have allowed differential interactions to be detected for different comparisons but data were too scarce for such an approach.

Conclusions

Our network meta-analysis suggests that patients with a formal diagnosis of PTSD and those with subclinical PTSD symptoms benefit from different psychological interventions. Those patients with a formal diagnosis, however, may benefit more from both specific and non-specific psychological interventions. We did not identify any intervention that was consistently superior to many or most other specific psychological interventions. Thus, we agree with the conclusion of Watts and colleagues that 'factors, such as access, acceptability and patient preference should exert strong and appropriate influence over the choice of treatment' (Watts *et al.* 2013, p. e547). Given the availability of effective treatment options and the severity of the disorder, the use of WL controls seems unethical and should be avoided in clinical trials. The effectiveness of EMDR and those interventions summarized in the OPI and ST categories in our analyses seems promising, but robust evidence from large trials with high study quality is lacking to date. In the future, large-sized trials should be conducted that compare such promising interventions with CBT and ET with robust evidence to expand available treatment options for PTSD.

Supplementary material

For supplementary material accompanying this paper visit http://dx.doi.org/10.1017/S0033291714000853.

Acknowledgements

We thank T. Barmettler for help with the literature search and trial retrieval, A. Volz, L. Trösch, R. Zurkinden and T. Tonia for help in the screening process, and H. Schmidt, J. Kummer and K. Abel for help with data extraction. We also thank K. Tal for providing editorial assistance with the manuscript.

J.B. and P.J. received a grant (no. 105314-118312/1) for this work from the Swiss National Science Foundation.

Declaration of Interest

None.

References

Altman DG, Bland JM (2003). Interaction revisited: the difference between two estimates. *British Medical Journal* **326**, 219.

Barth J, Munder T, Gerger H, Nüesch E, Trelle S, Znoj H, Jüni P, Cuijpers P (2013). Comparative efficacy of seven psychotherapeutic interventions for patients with depression: a network meta-analysis. *PLoS Medicine* **10**, e1001454.

Benish SG, Imel ZE, Wampold BE (2008). The relative efficacy of bona fide psychotherapies for treating post-traumatic stress disorder: a meta-analysis of direct comparisons. *Clinical Psychology Review* **28**, 746–758.

Bernat JA, Ronfeldt HM, Calhoun KS, Arias I (1998). Prevalence of traumatic events and peritraumatic predictors of posttraumatic stress symptoms in a nonclinical sample of college students. *Journal of Traumatic Stress* **11**, 645–664.

Bisson JI, Andrew M (2007). Psychological treatment of post-traumatic stress disorder (PTSD). *Cochrane Database of Systematic Reviews*. Issue 3, Art. No.: CD003388.

Bradley R, Greene J, Russ E, Dutra L, Westen D (2005). A multidimensional meta-analysis of psychotherapy for PTSD. *American Journal of Psychiatry* **162**, 214–227.

Breslau N (2009). The epidemiology of trauma, PTSD, and other posttrauma disorders. *Trauma, Violence and Abuse* **10**, 198–210.

Breslau N, Davis GC, Andreski P, Peterson E (1991). Traumatic events and posttraumatic stress disorder in an urban population of young adults. *Archives of General Psychiatry* **48**, 216–222.

Breslau N, Kessler RC, Chilcoat HD, Schultz LR, Davis GC, Andreski P (1998). Trauma and posttraumatic stress disorder in the community: the 1996 Detroit Area Survey of Trauma. *Archives of General Psychiatry* **55**, 626–632.

Cipriani A, Barbui C, Salanti G, Rendell J, Brown R, Stockton S, Purgato M, Spineli LM, Goodwin GM, Geddes JR (2011). Comparative efficacy and acceptability of antimanic drugs in acute mania: a multiple-treatments meta-analysis. *Lancet* **378**, 1306–1315.

Cipriani A, Furukawa TA, Salanti G, Geddes JR, Higgins J, Churchill R, Watanabe N, Nakagawa A, Omori IM, McGuire H, Tansella M, Barbui C (2009). Comparative efficacy and acceptability of 12 new-generation antidepressants: a multiple-treatments meta-analysis. *Lancet* **373**, 746–758.

Cloitre M (2009). Effective psychotherapies for posttraumatic stress disorder: a review and critique. *CNS Spectrums* **14**, 32–43.

Cohen J (1988). *Statistical Power Analysis for the Behavioral Sciences*, 2nd edn. Lawrence Erlbaum Associates: Hillsdale, NJ.

Cooper NJ, Sutton AJ, Lu G, Khunti K (2006). Mixed comparison of stroke prevention treatments in individuals with nonrheumatic atrial fibrillation. *Archives of Internal Medicine* **166**, 1269–1275.

Cuijpers P, Li J, Hofmann SG, Andersson G (2010*a*). Self-reported versus clinician-rated symptoms of depression as outcome measures in psychotherapy research on depression: a meta-analysis. *Clinical Psychology Review* **30**, 768–778.

Cuijpers P, Van Straten A, Bohlmeijer E, Hollon SD, Andersson G (2010*b*). The effects of psychotherapy for adult depression are overestimated: a meta-analysis of study quality and effect size. *Psychological Medicine* **40**, 211–223.

Davidson PR, Parker KC (2001). Eye movement desensitization and reprocessing (EMDR): a meta-analysis. *Journal of Consulting and Clinical Psychology* **69**, 305–316.

Egger M, Davey Smith G, Schneider M, Minder C (1997). Bias in meta-analysis detected by a simple, graphical test. *British Medical Journal* **315**, 629–634.

Egger M, Dickersin K, Davey Smith G (2001). *Problems and Limitations in Conducting Systematic Reviews*, 2nd edn. BMJ Books: London.

Egger M, Juni P, Bartlett C, Holenstein F, Sterne J (2003). How important are comprehensive literature searches and the assessment of trial quality in systematic reviews? Empirical study. *Health Technology Assessment* **7**, 1–4.

Ehlers A, Clark DM, Hackmann A, McManus F, Fennell M (2005). Cognitive therapy for post-traumatic stress disorder: development and evaluation. *Behaviour Research and Therapy* **43**, 413–431.

Foa EB, Rothbaum BO, Riggs DS, Murdock TB (1991). Treatment of posttraumatic stress disorder in rape victims: a comparison between cognitive-behavioral procedures and counseling. *Journal of Consulting and Clinical Psychology* **59**, 715–723.

Follmann D, Elliott P, Suh I, Cutler J (1992). Variance imputation for overviews of clinical trials with continuous response. *Journal of Clinical Epidemiology* **45**, 769–773.

Frost N, Laska KM, Wampold BE (2014). The evidence for present-centered therapy as a treatment for

posttraumatic stress disorder. *Journal of Traumatic Stress* **27**, 1–8.

Gerger H, Munder T, Barth J (2013). Specific and non-specific psychological interventions for PTSD symptoms: a meta-analysis with problem complexity as a moderator. *Journal of Clinical Psychology*. Published online: 18 December 2013. doi: 10.1002/jclp.22059.

Jüni P, Altman DG, Egger M (2001). Systematic reviews in health care: assessing the quality of controlled clinical trials. *British Medical Journal* **323**, 42–46.

Kessler RC, Sonnega A, Bromet E, Hughes M, Nelson CB (1995). Posttraumatic stress disorder in the National Comorbidity Survey. *Archives of General Psychiatry* **52**, 1048–1060.

Koren D, Arnon I, Klein E (2001). Long term course of chronic posttraumatic stress disorder in traffic accident victims: a three-year prospective follow-up study. *Behaviour Research and Therapy* **39**, 1449–1458.

Lu G, Ades AE (2004). Combination of direct and indirect evidence in mixed treatment comparisons. *Statistics in Medicine* **23**, 3105–3124.

Lumley T (2002). Network meta-analysis for indirect treatment comparisons. *Statistics in Medicine* **21**, 2313–2324.

Marshall RD, Olfson M, Hellman F, Blanco C, Guardino M, Struening EL (2001). Comorbidity, impairment, and suicidality in subthreshold PTSD. *American Journal of Psychiatry* **158**, 1467–1473.

Matt GE, Cook TD (2009). Threats to the validity of generalized inferences. In *The Handbook of Research Synthesis and Meta-Analysis*, 2nd edn (ed. H. Cooper, L. V. Hedges and J. C. Valentine), pp. 537–560. Russell Sage: New York.

Mayou RA, Ehlers A, Bryant B (2002). Posttraumatic stress disorder after motor vehicle accidents: 3-year follow-up of a prospective longitudinal study. *Behaviour Research and Therapy* **40**, 665–675.

Mendes DD, Mello MF, Ventura P, de Medeiros Passarela C, de Jesus Mari J (2008). A systematic review on the effectiveness of cognitive behavioral therapy for posttraumatic stress disorder. *International Journal of Psychiatry in Medicine* **38**, 241–259.

Munder T, Flückiger C, Gerger H, Wampold BE, Barth J (2012). Is the allegiance effect an epiphenomenon of true efficacy differences between treatments? A meta-analysis. *Journal of Counseling Psychology* **59**, 631–637.

Norris FH (1992). Epidemiology of trauma: frequency and impact of different potentially traumatic events on different demographic groups. *Journal of Consulting and Clinical Psychology* **60**, 409–418.

Nüesch E, Trelle S, Reichenbach S, Rutjes AWS, Bürgi E, Scherer M, Altman DG, Jüni P (2009). The effects of excluding patients from the analysis in randomised controlled trials: meta-epidemiological study. *British Medical Journal* **339**, b3244.

Nüesch E, Trelle S, Reichenbach S, Rutjes AWS, Tschannen B, Altman DG, Egger M, Jüni P (2010). Small study effects in meta-analyses of osteoarthritis trials: meta-epidemiological study. *British Medical Journal* **341**, c3515.

Perkonigg A, Pfister H, Stein MB, Höfler M, Lieb R, Maerker A, Wittchen HU (2005). Longitudinal course of posttraumatic stress disorder and posttraumatic stress disorder symptoms in a community sample of adolescents and young adults. *American Journal of Psychiatry* **162**, 1320–1327.

Peters JL, Sutton AJ, Jones DR, Abrams KR, Rushton L (2008). Contour-enhanced meta-analysis funnel plots help distinguish publication bias from other causes of asymmetry. *Journal of Clinical Epidemiology* **61**, 991–996.

Powers MB, Halpern JM, Ferenschak MP, Gillihan SJ, Foa EB (2010). A meta-analytic review of prolonged exposure for posttraumatic stress disorder. *Clinical Psychology Review* **30**, 635–641.

Reichenbach S, Sterchi R, Scherer M, Trelle S, Bürgi E, Bürgi U, Dieppe PA, Jüni P (2007). Meta-analysis: chondroitin for osteoarthritis of the knee or hip. *Annals of Internal Medicine* **146**, 580–590.

Resick PA, Galovski TE, Uhlmansiek MOB, Scher CD, Clum GA, Young-Xu Y (2008). A randomized clinical trial to dismantle components of cognitive processing therapy for posttraumatic stress disorder in female victims of interpersonal violence. *Journal of Consulting and Clinical Psychology* **76**, 243–258.

Resnick HS, Kilpatrick DG, Dansky BS, Saunders BE, Best CL (1993). Prevalence of civilian trauma and posttraumatic stress disorder in a representative national sample of women. *Journal of Consulting and Clinical Psychology* **61**, 984–991.

Rücker G, Schwarzer G, Carpenter JR, Schumacher M (2008). Undue reliance on I^2 in assessing heterogeneity may mislead. *BMC Medical Research Methodology* **8**, 79.

Salanti G, Higgins JPT, Ades AE, Ioannidis JPA (2008). Evaluation of networks of randomized trials. *Statistical Methods in Medical Research* **17**, 279–301.

Seidler G, Wagner F (2006). Comparing the efficacy of EMDR and trauma-focused cognitive-behavioral therapy in the treatment of PTSD: a meta-analytic study. *Psychological Medicine* **36**, 1515–1522.

Smith TC, Spiegelhalter DJ, Thomas A (1995). Bayesian approaches to random-effects meta-analysis: a comparative study. *Statistics in Medicine* **14**, 2685–2699.

Spiegelhalter DJ, Best NG, Carlin BP, Van Der Linde A (2002). Bayesian measures of model complexity and fit. *Journal of the Royal Statistical Society: Series B (Statistical Methodology)* **64**, 583–639.

Stein MB, Walker JR, Hazen AL, Forde DR (1997). Full and partial posttraumatic stress disorder: findings from a community survey. *American Journal of Psychiatry* **154**, 1114–1119.

Thompson SG, Higgins JPT (2002). How should meta-regression analyses be undertaken and interpreted? *Statistics in Medicine* **21**, 1559–1573.

van Etten ML, Taylor S (1998). Comparative efficacy of treatments for post-traumatic stress disorder: a meta-analysis. *Clinical Psychology and Psychotherapy* **5**, 126–144.

Walker EA, Gelfand A, Katon WJ, Koss MP, Von Korff M, Bernstein D, Russo J (1999). Adult health status of women

with histories of childhood abuse and neglect. *American Journal of Medicine* **107**, 332–339.

Wampold BE, Imel ZE, Laska KM, Benish S, Miller SD, Flückiger C, Del Re AC, Baardseth TP, Budge S (2010). Determining what works in the treatment of PTSD. *Clinical Psychology Review* **30**, 923–933.

Watts BV, Schnurr PP, Mayo L, Young-Xu Y, Weeks WB, Friedman MJ (2013). Meta-analysis of the efficacy of treatments for posttraumatic stress disorder. *Journal of Clinical Psychiatry* **74**, e541–e550.

Wood L, Egger M, Gluud LL, Schulz KF, Jüni P, Altman DG, Gluud C, Martin RM, Wood AJ, Sterne JA (2008). Empirical evidence of bias in treatment effect estimates in controlled trials with different interventions and outcomes: meta-epidemiological study. *British Medical Journal* **336**, 601–605.

Zlotnick C, Franklin CL, Zimmerman M (2002). Does 'subthreshold' posttraumatic stress disorder have any clinical relevance? *Comprehensive Psychiatry* **43**, 413–419.